

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/194,164 A
Source: IFW16
Date Processed by STIC: 08/25/2005

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 08/25/2005

PATENT APPLICATION: US/09/194,164A

TIME: 16:07:29

Input Set : A:\31608-20001.21.txt

Output Set: N:\CRF4\08252005\I194164A.raw

SEQUENCE LISTING

THE

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Dan, Michael D.

7 Kaplan, Howard A.

8 Maiti, Pradipto K.

11 (ii) TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS THAT

12 SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING

13 FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND

14 DETECTION OF CANCERS

16 (iii) NUMBER OF SEQUENCES: 30

18 (iv) CORRESPONDENCE ADDRESS:

19 (A) ADDRESSEE: Morrison & Foerster LLP

20 (B) STREET: 755 Page Mill Road

21 (C) CITY: Palo Alto

22 (D) STATE: CA

23 (E) COUNTRY: USA

24 (F) ZIP: 94304-1018

26 (v) COMPUTER READABLE FORM:

27 (A) MEDIUM TYPE: Floppy disk

28 (B) COMPUTER: IBM PC compatible

29 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

30 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

32 (vi) CURRENT APPLICATION DATA:

C--> 33 (A) APPLICATION NUMBER: US/09/194,164A

C--> 34 (B) FILING DATE: 09-Apr-1999

W--> 45 (C) CLASSIFICATION:

42 (vii) PRIOR APPLICATION DATA:

38 (A) APPLICATION NUMBER: PCT/US97/08962

39 (B) FILING DATE: 22-MAY-1997

43 (A) APPLICATION NUMBER: US 08/657,449

44 (B) FILING DATE: 22-MAY-1996

47 (viii) ATTORNEY/AGENT INFORMATION:

48 (A) NAME: Glaister, Debra J.

49 (B) REGISTRATION NUMBER: 33,888

50 (C) REFERENCE/DOCKET NUMBER: 31608-20001.21

52 (ix) TELECOMMUNICATION INFORMATION:

53 (A) TELEPHONE: (650) 813-5600

54 (B) TELEFAX: (650) 494-0792

58 (2) INFORMATION FOR SEQ ID NO: 1:

60 (i) SEQUENCE CHARACTERISTICS:

61 (A) LENGTH: 543 base pairs

62 (B) TYPE: nucleic acid

63 (C) STRANDEDNESS: double

RAW SEQUENCE LISTING

DATE: 08/25/2005

PATENT APPLICATION: US/09/194,164A

TIME: 16:07:29

Input Set : A:\31608-20001.21.txt

Output Set: N:\CRF4\08252005\I194164A.raw

```

64         (D) TOPOLOGY: linear
68     (ix) FEATURE:
69         (A) NAME/KEY: CDS
70         (B) LOCATION: 1..543
73     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
75 CAAGCTATTT AGGTGACACT ATAGAATACT CAAGCTATGC ATCCAACGCG TTGGGAGCTC      60
77 TCCCATATGG TCGACCTGCA GCGGCCGCA CTAGTGATTT CAAGCTTCAT CACTGAACAC      120
79 AGAGGACTCA CCATGGAGTT TGGGCTGAGC TGGGTTTTC TCGTTGCTCT TTTAAGAGGT      180
81 ATCCAGTGTC AGGTGCAGCT GGTGGAGTCT GGGGGAGGCG TGGTCCAGCC TGGGAGGTCC      240
83 CTGAGACTCT CCTGTGCAGC CTCTGGATTG CCCTTCAGAA GCTTTGCTAT GCACTGGGTC      300
85 CGCCAGGCTC TAGGCAAGGG GCTGGAGTGG GTGGCAGTTA TATCATATGA TGAAGCACT      360
87 AAATACTACG CAGACTCCGT GAAGGGGCGA TTCACCATCT CCAGAGACAC TTCCAAGAAC      420
89 ACGGTGTATC TAAAAATGAA CAGGCTGAGA ACTGAGGACA CGGCTGTCTT TTAATTGTGC      480
91 GAAAGACAGA GCCTGCTGGG TGACTATGAC CACTACTACG GNTTGGACGC TTGGGGAAAG      540
93 GGA
95 (2) INFORMATION FOR SEQ ID NO: 2:
97     (i) SEQUENCE CHARACTERISTICS:
98         (A) LENGTH: 179 amino acids
99         (B) TYPE: amino acid
100        (C) STRANDEDNESS: single
101        (D) TOPOLOGY: linear
107     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
109     Gln Ala Ile Val Thr Leu Asn Thr Gln Ala Met His Pro Thr Arg Trp
110     1          5          10          15
112     Glu Leu Ser His Met Val Asp Leu Gln Ala Ala Ala Leu Val Ile Ser
113     20          25          30
115     Ser Phe Ile Thr Glu His Arg Gly Leu Thr Met Glu Phe Gly Leu Ser
116     35          40          45
118     Trp Val Phe Leu Val Ala Leu Arg Gly Ile Gln Cys Gln Val Gln
119     50          55          60
121     Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg Ser Leu Arg
122     65          70          75          80
124     Leu Ser Cys Ala Ala Ser Gly Phe Pro Phe Arg Ser Phe Ala Met His
125     85          90          95
127     Trp Val Arg Gln Ala Leu Gly Lys Gly Leu Glu Trp Val Ala Val Ile
128     100         105         110
130     Ser Tyr Asp Gly Ser Thr Lys Tyr Tyr Ala Asp Ser Val Lys Gly Arg
131     115         120         125
133     Phe Thr Ile Ser Arg Asp Thr Ser Lys Asn Thr Val Tyr Leu Lys Met
134     130         135         140
136     Asn Arg Leu Arg Thr Glu Asp Thr Ala Val Phe Tyr Leu Cys Glu Arg
137     145         150         155         160
139     Gln Ser Leu Leu Gly Asp Tyr Asp His Tyr Tyr Gly Leu Asp Ala Trp
140     165         170         175
142     Gly Lys Gly
145 (2) INFORMATION FOR SEQ ID NO: 3:
147     (i) SEQUENCE CHARACTERISTICS:
148         (A) LENGTH: 543 base pairs
149         (B) TYPE: nucleic acid

```

RAW SEQUENCE LISTING

DATE: 08/25/2005

PATENT APPLICATION: US/09/194,164A

TIME: 16:07:29

Input Set : A:\31608-20001.21.txt

Output Set: N:\CRF4\08252005\I194164A.raw

```

150      (C) STRANDEDNESS: double
151      (D) TOPOLOGY: linear
157      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
159 TCCCTTTCCC CAAGCGTCCA ANCCGTAGTA GTGGTCATAG TCACCCAGCA GGCTCTGTCT      60
161 TTCGCACAAG TAAAAGACAG CCGTGTCCTC AGTTCTCAGC CTGTTTCATTT TTAGATACAC      120
163 CGTGTTCTTG GAAGTGTCTC TGGAGATGGT GAATCGCCCC TTCACGGAGT CTGCGTAGTA      180
165 TTTAGTGCTT CCATCATATG ATATAACTGC CACCCACTCC AGCCCCCTTGC CTAGAGCCTG      240
167 GCGGACCCAG TGCATAGCAA AGCTTCTGAA GGGGAATCCA GAGGCTGCAC AGGAGAGTCT      300
169 CAGGGACCTC CCAGGCTGGA CCACGCCTCC CCCAGACTCC ACCAGCTGCA CCTGACACTG      360
171 GATACCTCTT AAAAGAGCAA CGAGGAAAAC CCAGCTCAGC CCAAACCTCCA TGGTGAGTCC      420
173 TCTGTGTTCA GTGATGAAGC TTGAAATCAC TAGTGCGGCC GCCTGCAGGT CGACCATATG      480
175 GGAGAGCTCC CAACGCGTTG GATGCATAGC TTGAGTATTC TATAGTGTC AATAAATAGC      540
177 TTG                                         543

179 (2) INFORMATION FOR SEQ ID NO: 4:
181      (i) SEQUENCE CHARACTERISTICS:
182          (A) LENGTH: 450 base pairs
183          (B) TYPE: nucleic acid
184          (C) STRANDEDNESS: double
185          (D) TOPOLOGY: linear
189      (ix) FEATURE:
190          (A) NAME/KEY: CDS
191          (B) LOCATION: 1..450
194      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
196 CTCGAGATGG ACATGGAGTT CCAGGCGCAG CTTCTCTTCC TCCTGCTACT CTGGCTCCCA      60
198 GATATCACCG GAGATATTGT GTTGACGCAG TCTCCAGGCA CCCTGTCTTT GTCTCCAGGG      120
200 GAAAGAGCCA CCCTCTCCTG CAGGGCCAGT CAGAGTGTTA GTAGCAGCTA CTTAGCCTGG      180
202 TACCAGCAGA AACCTGGCCA GGCTCCCAGG CTCCTCATCT ATGGTGCATC CACCAGGGCC      240
204 ACTGGCATGC CAGACAGGTC CAGTGGCAGT GGGTCCGGGA CAGACTTCAC TCTCACCATC      300
206 AGTAGACTGG AGCCTGAAGA TTTTGCAGTG TATTACTGTC AGCAGTATGG TAGCTCACCT      360
208 CAGACACCTC AGATCACTTT CGGCGGAGGG ACCAAGGTGG AGATCAAACG AACTGTGGCT      420
210 GCACCATCTG TCTTCATCTT CCCGCCATCT                                         450

212 (2) INFORMATION FOR SEQ ID NO: 5:
214      (i) SEQUENCE CHARACTERISTICS:
215          (A) LENGTH: 150 amino acids
216          (B) TYPE: amino acid
217          (C) STRANDEDNESS: single
218          (D) TOPOLOGY: linear
224      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
226 Leu Glu Met Asp Met Glu Phe Gln Ala Gln Leu Leu Phe Leu Leu Leu
227 1           5           10           15
229 Leu Trp Leu Pro Asp Ile Thr Gly Asp Ile Val Leu Thr Gln Ser Pro
230           20           25           30
232 Gly Thr Leu Ser Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg
233           35           40           45
235 Ala Ser Gln Ser Val Ser Ser Ser Tyr Leu Ala Trp Tyr Gln Gln Lys
236           50           55           60
238 Pro Gly Gln Ala Pro Arg Leu Leu Ile Tyr Gly Ala Ser Thr Arg Ala
239           65           70           75           80
241 Thr Gly Met Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe

```

RAW SEQUENCE LISTING

DATE: 08/25/2005

PATENT APPLICATION: US/09/194,164A

TIME: 16:07:29

Input Set : A:\31608-20001.21.txt

Output Set: N:\CRF4\08252005\I194164A.raw

```

242          85          90          95
244    Thr Leu Thr Ile Ser Arg Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr
245          100          105          110
247    Cys Gln Gln Tyr Gly Ser Ser Pro Gln Thr Pro Gln Ile Thr Phe Gly
248          115          120          125
250    Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val
251          130          135          140
253    Phe Ile Phe Pro Pro Ser
254          145          150
256 (2) INFORMATION FOR SEQ ID NO: 6:
258     (i) SEQUENCE CHARACTERISTICS:
259         (A) LENGTH: 450 base pairs
260         (B) TYPE: nucleic acid
261         (C) STRANDEDNESS: double
262         (D) TOPOLOGY: linear
268     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
270 AGATGGCGGG AAGATGAAGA CAGATGGTGC AGCCACAGTT CGTTTGATCT CCACCTTGGT      60
272 CCCTCCGCCG AAAGTGATCT GAGGTGTCTG AGGTGAGCTA CCATACTGCT GACAGTAATA     120
274 CACTGCAAAA TCTTCAGGCT CCAGTCTACT GATGGTGAGA GTGAAGTCTG TCCCGGACCC     180
276 ACTGCCACTG AACCTGTCTG GCATGCCAGT GGCCCTGGTG GATGCACCAT AGATGAGGAG     240
278 CCTGGGAGCC TGGCCAGGTT TCTGCTGGTA CCAGGCTAAG TAGCTGCTAC TAACACTCTG     300
280 ACTGGCCCTG CAGGAGAGGG TGGCTCTTTC CCCTGGAGAC AAAGACAGGG TGCCTGGAGA     360
282 CTGCGTCAAC ACAATATCTC CGGTGATATC TGGGAGCCAG AGTAGCAGGA GGAAGAGAAG     420
284 CTGCGCCTGG AACTCCATGT CCATCTCGAG                                     450
286 (2) INFORMATION FOR SEQ ID NO: 7:
288     (i) SEQUENCE CHARACTERISTICS:
289         (A) LENGTH: 34 base pairs
290         (B) TYPE: nucleic acid
291         (C) STRANDEDNESS: single
292         (D) TOPOLOGY: linear
298     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
300 TATGAAGACA CCAGGCCGAT ATTGTGTTGA CGCA                                     34
302 (2) INFORMATION FOR SEQ ID NO: 8:
304     (i) SEQUENCE CHARACTERISTICS:
305         (A) LENGTH: 26 base pairs
306         (B) TYPE: nucleic acid
307         (C) STRANDEDNESS: single
308         (D) TOPOLOGY: linear
314     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
316 TATCCGGATG CAGCCACAGT TCGTTT                                     26
318 (2) INFORMATION FOR SEQ ID NO: 9:
320     (i) SEQUENCE CHARACTERISTICS:
321         (A) LENGTH: 26 base pairs
322         (B) TYPE: nucleic acid
323         (C) STRANDEDNESS: single
324         (D) TOPOLOGY: linear
330     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
332 TATTCGGACA GGTGCAGCTG GTGGAG                                     26
334 (2) INFORMATION FOR SEQ ID NO: 10:

```

RAW SEQUENCE LISTING

DATE: 08/25/2005

PATENT APPLICATION: US/09/194,164A

TIME: 16:07:29

Input Set : A:\31608-20001.21.txt

Output Set: N:\CRF4\08252005\I194164A.raw

```

336      (i) SEQUENCE CHARACTERISTICS:
337          (A) LENGTH: 27 base pairs
338          (B) TYPE: nucleic acid
339          (C) STRANDEDNESS: single
340          (D) TOPOLOGY: linear
346      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
348 TATGGATCCT GAGGAGACGG TGACCGT                                27
350 (2) INFORMATION FOR SEQ ID NO: 11:
352      (i) SEQUENCE CHARACTERISTICS:
353          (A) LENGTH: 60 base pairs
354          (B) TYPE: nucleic acid
355          (C) STRANDEDNESS: single
356          (D) TOPOLOGY: linear
362      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
364 TATATATCCG GAGGTGGTGG ATCAGGTGGA GGTGGCTCCC AGGTGCAGCT GGTGGAGTCT    60
367 (2) INFORMATION FOR SEQ ID NO: 12:
369      (i) SEQUENCE CHARACTERISTICS:
370          (A) LENGTH: 46 base pairs
371          (B) TYPE: nucleic acid
372          (C) STRANDEDNESS: single
373          (D) TOPOLOGY: linear
379      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
381 ACCTCCGGAA CCGCCACCGC CAGAGACAGA TGGTGCAGCC ACATTC                    46
383 (2) INFORMATION FOR SEQ ID NO: 13:
385      (i) SEQUENCE CHARACTERISTICS:
386          (A) LENGTH: 918 base pairs
387          (B) TYPE: nucleic acid
388          (C) STRANDEDNESS: single
389          (D) TOPOLOGY: linear
393      (ix) FEATURE:
394          (A) NAME/KEY: CDS
395          (B) LOCATION: join(1..906, 913..918)
398      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
400 GAA TTC ATG AAA AAA ACC GCT ATC GCG ATC GCA GTT GCA CTG GCT GGT      48
401 Glu Phe Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly
402   1           5           10           15
404 TTC GCT ACC GTT GCG CAG GCC GAT ATT GTG TTG ACG CAG TCT CCA GGC      96
405 Phe Ala Thr Val Ala Gln Ala Asp Ile Val Leu Thr Gln Ser Pro Gly
406           20           25           30
408 ACC CTG TCT TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC AGG GCC     144
409 Thr Leu Ser Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala
410           35           40           45
412 AGT CAG AGT GTT AGT AGC AGC TAC TTA GCC TGG TAC CAG CAG AAA CCT     192
413 Ser Gln Ser Val Ser Ser Ser Tyr Leu Ala Trp Tyr Gln Gln Lys Pro
414           50           55           60
416 GGC CAG GCT CCC AGG CTC CTC ATC TAT GGT GCA TCC ACC AGG GCC ACT     240
417 Gly Gln Ala Pro Arg Leu Leu Ile Tyr Gly Ala Ser Thr Arg Ala Thr
418           65           70           75           80
420 GGC ATG CCA GAC AGG TTC AGT GGC AGT GGG TCC GGG ACA GAC TTC ACT     288

```

VERIFICATION SUMMARY

DATE: 08/25/2005

PATENT APPLICATION: US/09/194,164A

TIME: 16:07:30

Input Set : A:\31608-20001.21.txt

Output Set: N:\CRF4\08252005\I194164A.raw

L:33 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:34 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:40 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1)(vi)
L:45 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1)(vi)